



Published in final edited form as:

Osteoporos Int. 2011 April ; 22(4): 1123–1131. doi:10.1007/s00198-010-1290-6.

Reliability of vertebral fracture assessment using multidetector CT lateral scout views: the Framingham Osteoporosis Study

E. J. Samelson,

Hebrew SeniorLife, Institute for Aging Research Boston, Boston, MA, USA; Department of Medicine, Harvard Medical School, Boston, MA, USA

B. A. Christiansen,

Beth Israel Deaconess Medical Center, Center for Advanced Orthopedic Studies, Boston, MA, USA; Department of Orthopedic Surgery, Harvard Medical School, Boston, MA, USA; Department of Orthopaedics, University of California Davis, Sacramento, CA, USA
bchristiansen@ucdavis.edu

S. Demissie,

Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA
demissie@bu.edu

K. E. Broe,

Hebrew SeniorLife, Institute for Aging Research Boston, Boston, MA, USA
broe@hrca.harvard.edu

Y. Zhou,

Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA
zyanhua@bu.edu

C. A. Meng,

Hebrew SeniorLife, Institute for Aging Research Boston, Boston, MA, USA Ching-AnMeng@hrca.harvard.edu

W. Yu,

Radiology, Peking Union Medical College Hospital, Beijing, China weiyu5508@yahoo.com

X. Cheng,

Radiology, Beijing Ji Shui Tan hospital, Beijing, China xiao65@263.net

C. J. O'Donnell,

Department of Medicine, Harvard Medical School, Boston, MA, USA; National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA, USA codonnell@nih.gov

U. Hoffmann,

Department of Medicine, Harvard Medical School, Boston, MA, USA

H. K. Genant, D. P. Kiel, and

Hebrew SeniorLife, Institute for Aging Research Boston, Boston, MA, USA; Department of Medicine, Harvard Medical School, Boston, MA, USA kiel@hrca.harvard.edu

M. L. Bouxsein

Beth Israel Deaconess Medical Center, Center for Advanced Orthopedic Studies, Boston, MA, USA; Department of Orthopedic Surgery, Harvard Medical School, Boston, MA, USA
mbouxsei@bidmc.harvard.edu

Abstract

Summary—Two radiologists evaluated images of the spine from computed tomography (CT) scans on two occasions to diagnose vertebral fracture in 100 individuals. Agreement was fair to good for mild fractures, and agreement was good to excellent for more severe fractures. CT scout views are useful to assess vertebral fracture.

Introduction—We investigated inter-reader agreement between two radiologists and intra-reader agreement between duplicate readings for each radiologist, in assessment of vertebral fracture using a semi-quantitative method from lateral scout views obtained by CT.

Methods—Participants included 50 women and 50 men (age 50-87 years, mean 70 years) in the Framingham Study. T4-L4 vertebrae were assessed independently by two radiologists on two occasions using a semi-quantitative scale as normal, mild, moderate, or severe fracture.

Results—Vertebra-specific prevalence of grade ≥ 1 (mild) fracture ranged from 3% to 5%. We found fair ($\kappa=56-59\%$) inter-reader agreement for grade ≥ 1 vertebral fractures and good ($\kappa=68-72\%$) inter-reader agreement for grade ≥ 2 fractures. Intra-reader agreement for grade ≥ 1 vertebral fracture was fair ($\kappa=55\%$) for one reader and excellent for another reader ($\kappa=77\%$), whereas intra-reader agreement for grade ≥ 2 vertebral fracture was excellent for both readers ($\kappa=76\%$ and 98%). Thoracic vertebrae were more difficult to evaluate than the lumbar region, and agreement was lowest (inter-reader $\kappa=43\%$) for fracture at the upper (T4-T9) thoracic levels and highest (inter-reader $\kappa=76-78\%$) for the lumbar spine (L1-L4).

Conclusions—Based on a semi-quantitative method to classify vertebral fractures using CT scout views, agreement within and between readers was fair to good, with the greatest source of variation occurring for fractures of mild severity and for the upper thoracic region. Agreement was good to excellent for fractures of at least moderate severity. Lateral CT scout views can be useful in clinical research settings to assess vertebral fracture.

Keywords

Computed tomography; Lateral scout; Reliability; Scoutviews; Semiquantitative; Vertebral fracture

Introduction

Vertebral fractures are the hallmark of osteoporosis and the most common osteoporotic fracture affecting as many as half of US white women and men in their 80s [1–3]. In Europe, one in eight individuals greater than 50 years have vertebral fractures [4], and worldwide, the estimated number of clinical vertebral fractures for the year 2000 was 1.4 million [5]. In 2005, incident vertebral fractures were responsible for more than 1.0 billion dollars in direct healthcare costs in the US [6]. Vertebral fractures cause pain, disfigurement, depression, and functional impairment [7–9]. Moreover, vertebral fractures are harbingers for future disabling fractures as well as death [10–12]. Despite the impact of vertebral fracture, less than one third of individuals with vertebral fractures receive medical attention and even fewer are treated [13–15].

Numerous methods have been developed to assess vertebral fracture from conventional lateral radiographs [16,17]. More recently, these methods for detecting vertebral fracture have been applied to lateral spine images from dual-energy X-ray absorptiometry (DXA)

vertebral fracture assessment (VFA) [18–20] with the intention to improve diagnosis and treatment without increasing cost and exposure to radiation. Increasingly, older individuals are undergoing thoracoabdominal computed tomography (CT) examinations that include acquisition of posterior-anterior and lateral scout view images which are 2D digital radiographs used to localize the area to be imaged. Although resolution of a scout view is lower than for a conventional radiograph, there is increasing interest in use of CT scout view images for evaluation of vertebral fracture [21,22]. Whereas reliability of vertebral fracture assessment using radiographs and VFA has been well studied [16,17,20,23–28], there is only a single prior study [22], conducted more than a decade ago, that reported reliability of vertebral fracture assessment from CT lateral scout views. Since this study, there have been significant advances in CT technology, allowing for faster imaging and better resolution.

As part of a large observational study of osteoporosis, we are evaluating CT image data of the spine acquired for participants in the Framingham Study who underwent cardiac imaging studies. Our protocol calls for vertebral fracture assessment from lateral CT scout views using a semi-quantitative method, and we aimed to evaluate reliability of our measurements. Therefore, we conducted a study to determine inter- and intra-reader reliability for vertebral fracture assessment using lateral spine scout views obtained as part of CT examinations for a sample of 100 participants.

Methods

Subjects

Participants for the current study were selected from the Framingham Heart Study Offspring and Third Generation Multi-detector Computed Tomography (MDCT) Study [29]. Members of the Offspring and Third Generation Cohorts include second and Third Generation Offspring (and their spouses) of the original cohort that was established in 1948 [30–32]. In 2002–2005, QCT scans of the chest and abdomen were acquired in 3,529 participants in the MDCT Study (age 40–90 years, mean 51 years) for assessment of coronary and aortic calcium [33]. Exclusion criteria for the MDCT Study included pregnancy or age less than 40 years for women, age less than 35 years for men, and weight greater than 320 pounds.

We selected 100 subjects to conduct a reliability study of vertebral fracture identification. To ensure an adequate number of individuals with vertebral fractures in our reliability study, a clinical investigator experienced in evaluating lateral spine images reviewed the scout images of persons age 70 years and older to identify 16 individuals (eight women and eight men) with suspected vertebral fracture [34]. We then selected a convenience sample of 84 additional persons so that our study group included 100 individuals (50 women and 50 men) in the age groups representing the distribution of age in the parent study and giving more weight to the older groups with greater prevalence of fracture, as follows: seven persons age 50–54 years and seven persons age 55–59 years, 14 persons age 60–64 years, 16 persons age 65–69 years, 19 persons 70–74 years, and 37, age 75+ years.

Computed tomography scan

An 8-slice multidetector computed tomography scanner (Lightspeed Ultra, General Electric Medical Systems, Milwaukee, WI, USA) was used for cardiac imaging of calcification [29,35]. The scout views consisted of frontal and lateral low energy 2D scanograms extending from the upper thoracic (T4) to sacral (S1) vertebral levels.

Study design and assessment of prevalent vertebral fracture

The lateral scout images for the 100 subjects were evaluated twice (“Time 1” and “Time 2”), several weeks apart, by each of two independent readers (“Reader A” and “Reader B”).

Thus, 4 evaluations were independently performed for each subject's scout view image. The scans for the 100 subjects in the reliability study were randomly ordered within five larger sets of scans for evaluation in the parent study. Details of the reliability study were not revealed to readers and images were stripped of demographic and any other identifying information. In addition, we provided different identification numbers associated with the 100 subjects for each reading set so that each batch appeared to the readers as a unique set of subjects. Also, the order of the subject scans was changed for each evaluation.

The two readers were experienced radiologists who underwent training with a musculoskeletal radiologist who developed a semi-quantitative method to assess vertebral fracture [34]. Prior to this reliability study, CT scout views for 45 subjects randomly selected from the parent study were used as a training set for readers along with a radiographic atlas. We provided readers with an automated system and instruction manual to directly enter findings into an electronic database.

Readers evaluated 13 vertebral levels from T4 to L4 and visually classified vertebrae as normal (grade 0), mildly deformed (grade 1, ~20–25% reduction in anterior, middle, and/or posterior height and a reduction in area 10–20%), moderately deformed (grade 2, ~25–40% reduction in any height and a reduction in area 20–40%), or severely deformed (grade 3, ~40% reduction in any height and area) [34]. Scan quality was recorded as “good” or “poor”. Vertebral levels unable to be adequately visualized were classified as ‘unreadable’. Vertebral deformities due to causes other than fracture, such as spondylosis and Scheuermann's disease, were classified as non-osteoporotic deformities. These were considered to be non-fractured in the analysis [36].

Prevalent vertebral fracture was defined as a vertebral body graded at least mildly deformed (grades 1–3). We also used a more restrictive classification of vertebral fracture, defined as moderately to severely deformed (at least ~25–40% reduction in any vertebral height or grades 2–3).

Statistical analysis

We calculated the frequency of unreadable vertebrae by spinal location for each reader at each of the two time points. Prevalence of fracture, per person and per vertebra, was calculated using two dichotomized definitions of fracture: [1] grades 1-3 (mild, moderate, and severe) versus grade 0 (normal), and [2] grades 2-3 (moderate and severe) versus grades 0-1 (normal and mild). In calculating prevalence, vertebrae that were unreadable were classified as normal.

The number and severity of fractured vertebrae is given for each of the two readers at each of the two time points. To calculate severity of vertebral fracture per person, individuals with more than one fracture were classified according to his/her most severe fracture grade (normal, mild, moderate, and severe).

We estimated agreement for fracture prevalence per vertebra, corrected for chance, using a simple kappa statistic and associated 95% confidence intervals (CI) [37]. We provide kappa for intra-reader agreement of fracture prevalence per vertebra between time 1 and time 2, separately for each of the two readers, and kappa for inter-reader agreement of fracture prevalence per vertebra between reader A and reader B, separately for each of the two time points.

We conducted stratified analysis to evaluate potential differences in agreement of fracture prevalence per vertebra by spinal region, categorized as T4-T9, T10-T12, and L1-L4 so that we could compare results to others [22] and have adequate numbers of fractures within each

region for meaningful statistical analysis. We calculated kappa (κ) to estimate inter- and intra-reader agreement for fracture prevalence per vertebra beyond that expected by chance, and considered kappa >0.75 as excellent agreement, $0.40-0.75$ as fair to good agreement, and <0.40 as poor agreement beyond that expected by chance as characterized by Fleiss [38]. Finally, we repeated our analysis including only those scans indicated as good quality by readers.

Results

Mean age of participants was 70 years ($SD=9$) and ranged from 50 to 87 years. The study group of 50 women and 50 men included predominantly members of the Offspring Cohort ($n=92$) and few members of the younger, Third Generation Cohort ($n=8$). Mean weight was $70 (\pm 17)$ kg in women and $85 (\pm 14)$ kg in men. Mean body mass index was $28 (\pm 6)$ kg/m² (Table 1).

The number of unreadable vertebrae was greatest for the upper thoracic spine (Fig. 1). Reader A classified a greater number of vertebrae as unreadable compared to reader B. For example, according to reader A, the number of unreadable vertebrae was 112 (9%) at time 1, which was four times greater than the number according to reader B, who classified 30 (2%) vertebrae as unreadable.

Each reader identified more fractured (grades 1–3) vertebrae at time 2 than time 1 (Fig. 2). Total number of vertebrae classified as fractured by reader A was 40 at time 1 and 52 at time 2. Number of vertebrae classified as fractured by reader B was 64 at time 1 and 66 at time 2. A bimodal distribution of prevalent vertebral fracture was observed with peak frequencies occurring at (and near) T8 and T12–L1 (Fig. 2). For example, prevalence (per person) was 11-12% at T8 and 7-8% at T12 (for three of the four evaluations).

Reader A identified 23-24 persons (23-24%) with at least one fracture, whereas reader B classified 35 persons (35%) with fracture (Table 2). Both raters classified fewer individuals with 1 vertebral fracture (and more individuals with two fractures) at the second reading compared to the first reading. The number of participants with three or more fractures doubled to eight persons at time 2 from four persons at time 1 for reader A, although reader B identified 1 fewer person with three or more fractures at the second evaluation (nine versus eight persons).

Reader A identified fewer mild (grade 1) fractures according time 1 relative to time 2 and fewer mild fractures than both evaluations for reader B (Table 3). For example, the number of vertebrae judged by reader A as mild fractures increased from the first to second evaluations from 25/1300 (2%) to 38/1300 (3%), whereas reader B classified 41-44 (3%) vertebral levels as mild fractures. The number of persons identified with mild fractures was consistent between time 1 and time 2 evaluations for each reader. However, reader A judged fewer individuals with mild fracture (13%) than reader B (22-23%). Frequency of moderate to severe fractures was low, with 14-15 vertebrae classified as \geq grade 2 by reader A and 22-23 vertebrae classified as \geq grade 2 by reader B. Per person prevalence of moderate to severe fractures was 10-11% according to reader A and 12-13% for reader B.

Intra-reader agreement for per vertebra prevalence of mild fracture (grade ≥ 1) was fair to good for reader A, $\kappa=0.55$, 95% CI: 0.43-0.67, and excellent for reader B $\kappa=0.77$, 95% CI: 0.69-0.85 (Table 4). Agreement for per vertebra prevalence of mild fracture between readers was good, and agreement was similar at time 1 ($\kappa=0.56$, 95% CI: 0.44-0.68) and time 2 ($\kappa=0.59$, 95% CI: 0.48-0.70). When we considered moderate-severe vertebral fractures (grades ≥ 2), intra-reader agreement was excellent for both readers, 0.76 (95% CI: 0.58-0.93)

for reader A and 0.98 for reader B (95% CI: 0.93-1.00). Agreement between readers reached 68% at time 1 and 72% at time 2 for moderate-severe fracture.

Intra-reader and inter-reader agreement for per vertebra prevalence of fractures grade ≥ 1 increased with descending vertebral level for each of the four evaluations (Table 5). For example, inter-reader agreement between readers at time 2 was 0.43 for T4-T9, 0.72 for T10-T12, and 0.78 for L1-L4, and the pattern was similar for the other evaluations.

Restricting the analysis to include only scans indicated as good image quality had little or no improvement on agreement for either intra- or inter-reader agreement.

Discussion

We determined reliability of a semi-quantitative method to evaluate prevalent vertebral fracture from CT lateral scout views. Four evaluations (by two readers at two time points) were independently performed to classify 1,300 vertebra in 100 participants in a community-based cohort. We found fair ($\kappa=56-59\%$) inter-reader agreement for \geq grade 1 vertebral fracture and good ($\kappa=68-72\%$) inter-reader agreement for \geq grade 2 fracture. Intra-reader agreement for \geq grade 1 vertebral fracture was fair ($\kappa=55\%$) for one reader and excellent for another reader ($\kappa=77\%$), whereas intra-reader agreement for \geq grade 2 vertebral fracture was excellent for both readers ($\kappa=76\%$ and 98%). Upper thoracic vertebrae were more likely to be unreadable than the lower thoracic and lumbar regions, and agreement was lowest (inter-reader $\kappa=43\%$) for fracture at the upper (T4-T9) thoracic levels evaluated and highest (inter-reader $\kappa=76-78\%$) for the lumbar spine.

Takada and colleagues [22] used similar methods as ours and reported higher inter- and intra-reader reliability than estimates found in the current study. Agreement for \geq grade 1 fracture between readers (at time 1) was $\kappa=0.68$, compared to our inter-reader reliability estimates of 0.56 for time 1 and 0.59 for time 2, and intra-reader agreement was 0.79 and 0.87 compared to our estimates of 0.55 and 0.77. Prevalence of fracture was 40% lower in our study which may explain, at least in part, lower estimates of agreement. The subjects in the study of Takada et al. [22] were women with mean age 56 years, and more than half were part of a clinical trial that required T-scores at least as low as -2.0 . In contrast, participants in our study were mean age 70 years and included equal numbers of women and men. While information on fracture severity was not provided in the study of Takada et al. [22], higher frequency of more severe fractures is expected in this group which may also help explain the higher reliability estimates compared to our results. In fact, agreement between readers for \geq grade 2 fracture in our study was 68-72%, comparable to Takada's findings, and inter-reader agreement for \geq grade 2 fracture was 76-98%, somewhat higher than reported for \geq grade 1 fracture by Takada et al. [22]. However, prevalence of \geq grade 2 fracture was low (1-2%) in our population-based study.

We found frequency of unreadable vertebrae was highest and intra-reader agreement was lowest for the upper thoracic spine. This was consistent with Takada et al. [22] and others who used different imaging modalities [20,39,40] and fracture classification methods [24]. Reproducibility of readings in the upper thoracic spine may be worse than other regions due to poorer image quality, in part attributable to the projection of other structures (pulmonary hilus and scapula) on the vertebral body. The lower levels of the spine have larger vertebrae and less obliquity and obscuring lung markings than the upper vertebrae. In contrast, Olmez et al. [41] reported, in a study of 67 postmenopausal women, little difference in agreement by vertebral level, assessed using the Kleerekopper method [42].

Intra-reader agreement for semi-quantitative vertebral fracture assessment using CT scout images in our study is lower than estimates reported for lateral spine images using DXA or

conventional radiography. In one study that examined reliability of DXA for fracture assessment in 203 postmenopausal women (mean age 68 years), intra-reader agreement, as measured by kappa, was 64-77% [18]. Intra-reader agreement is highest using lateral radiograph of the spine, considered the gold standard for vertebral fracture diagnosis, with a kappa estimate given at 89% [34]. Although image quality is lower for CT scout views than conventional radiography, CT scout views offer some advantages. Specifically, CT fan-beam images do not suffer from parallax distortion present in cone-beam imaging geometry as in conventional radiographs. Since a lateral scout view is acquired as part of conducting a CT study for purposes other than identification of spinal fracture, assessment of vertebral fracture using this method exposes individuals to lower radiation levels than conventional radiography, while at the same time identifies candidates for treatment who might not have otherwise been evaluated.

This study is limited by the relatively small number of fractures; however, prevalence is lower for community-based populations as in the current study than for subjects recruited from clinical settings.

Another limitation is that we did not perform sagittal midline reformations to improve the sensitivity of fracture detection. The scope of this study was to evaluate in 100 subjects reader reliability of methods used to assess vertebral fracture in participants in the parent study, the Framingham Heart Study Offspring and Third Generation MDCT Study [29]. Thus, although sagittal spine images will identify vertebral fractures not evident on the transverse axial images [43–46], use of reformations was beyond the scope of the current study.

This study is also limited by the lower spatial resolution and higher photon noise of CT scout views compared to other image modalities intended for evaluation of vertebral fracture, such as conventional radiography. In contrast to radiographs, however, CT scout views are not affected by parallax distortion which can reduce accuracy in fracture diagnosis. Additionally, information on reliability of the method used in this study is needed since these images allow identification of fracture in patients undergoing CT examinations for other purposes without additional procedures or radiation exposure. Given the high risk of fracture (and death) in older adults with prevalent vertebral fracture, identification of vertebral fracture from CT scout views provides opportunity for diagnosis and treatment to prevent pain, loss of function, and medical costs associated with subsequent fracture.

In conclusion, in this community-based sample of women and men, agreement between readers was fair to good for assessment of vertebral fractures using lateral CT scout views, (56-59% for mild fractures and 68-72% for moderate to severe fractures). Agreement for mild to severe fracture within readers was good to excellent (55-77%), whereas within-reader agreement for moderate to severe fracture was excellent (76-98%). Evaluation of lateral CT scout views using a semi-quantitative method is a reliable approach to identifying prevalent vertebral fractures. Lateral CT scout views can be useful in clinical settings to assess vertebral fracture.

Acknowledgments

This work was supported by NIH R01AR053986, R01AR/AG041398, K01 AR053118, T32 AG023480, and by the National Heart, Lung, and Blood Institute (NHLBI) Framingham Heart Study (NIH/NHLBI Contract N01-HC-25195).

References

1. Bone health and osteoporosis: a report of the surgeon general. Rockville, MD: US Department of Health and Human Services, Office of the Surgeon General; 2004.
2. Melton LJ 3rd, Lane AW, Cooper C, Eastell R, O'Fallon WM, Riggs BL. Prevalence and incidence of vertebral deformities. *Osteoporos Int* 1993;3:113–119. [PubMed: 8481586]
3. Davies KM, Stegman MR, Heaney RP, Recker RR. Prevalence and severity of vertebral fracture: the Saunders County Bone Quality Study. *Osteoporos Int* 1996;6:160–165. [PubMed: 8704356]
4. Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. *Bull World Health Organ* 2003;81:646–656. [PubMed: 14710506]
5. Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporos Int* 2006;17:1726–1733. [PubMed: 16983459]
6. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005–2025. *J Bone Miner Res* 2007;22:465–475. [PubMed: 17144789]
7. Edmond SL, Kiel DP, Samelson EJ, Kelly-Hayes M, Felson DT. Vertebral deformity, back symptoms, and functional limitations among older women: the Framingham Study. *Osteoporos Int* 2005;16:1086–1095. [PubMed: 15726295]
8. Melton LJ 3rd. Adverse outcomes of osteoporotic fractures in the general population. *J Bone Miner Res* 2003;18:1139–1141. [PubMed: 12817771]
9. Oleksik A, Lips P, Dawson A, Minshall ME, Shen W, Cooper C, Kanis J. Health-related quality of life in postmenopausal women with low BMD with or without prevalent vertebral fractures. *J Bone Miner Res* 2000;15:1384–1392. [PubMed: 10893688]
10. Cauley JA, Hochberg MC, Lui LY, Palermo L, Ensrud KE, Hillier TA, Nevitt MC, Cummings SR. Long-term risk of incident vertebral fractures. *Jama* 2007;298:2761–2767. [PubMed: 18165669]
11. Lau E, Ong K, Kurtz S, Schmier J, Edidin A. Mortality following the diagnosis of a vertebral compression fracture in the Medicare population. *J Bone Joint Surg Am* 2008;90:1479–1486. [PubMed: 18594096]
12. Hasserijs R, Karlsson MK, Jonsson B, Redlund-Johnell I, Johnell O. Long-term morbidity and mortality after a clinically diagnosed vertebral fracture in the elderly—a 12- and 22-year follow-up of 257 patients. *Calcif Tissue Int* 2005;76:235–242. [PubMed: 15812579]
13. Freedman BA, Potter BK, Nesti LJ, Giuliani JR, Hampton C, Kuklo TR. Osteoporosis and vertebral compression fractures—continued missed opportunities. *Spine J* 2008;8:756–762. [PubMed: 18343730]
14. Simon M. Recognition of osteoporosis since 1997. *Am J Public Health* 2002;92:1707. [PubMed: 12406786]
15. Siris ES, Bilezikian JP, Rubin MR, Black DM, Bockman RS, Bone HG, Hochberg MC, McClung MR, Schnitzer TJ. Pins and plaster aren't enough: a call for the evaluation and treatment of patients with osteoporotic fractures. *J Clin Endocrinol Metab* 2003;88:3482–3486. [PubMed: 12915621]
16. Black DM, Palermo L, Nevitt MC, Genant HK, Epstein R, San Valentin R, Cummings SR. Comparison of methods for defining prevalent vertebral deformities: the study of osteoporotic fractures. *J Bone Miner Res* 1995;10:890–902. [PubMed: 7572313]
17. Ferrar L, Jiang G, Schousboe JT, DeBold CR, Eastell R. Algorithm-based qualitative and semiquantitative identification of prevalent vertebral fracture: agreement between different readers, imaging modalities, and diagnostic approaches. *J Bone Miner Res* 2008;23:417–424. [PubMed: 17967136]
18. Fuerst T, Wu C, Genant HK, von Ingersleben G, Chen Y, Johnston C, Econs MJ, Binkley N, Vokes TJ, Crans G, Mitlak BH. Evaluation of vertebral fracture assessment by dual X-ray absorptiometry in a multicenter setting. *Osteoporos Int* 2009;20:1199–1205. [PubMed: 19083074]
19. Damiano J, Kolta S, Porcher R, Tournoux C, Dougados M, Roux C. Diagnosis of vertebral fractures by vertebral fracture assessment. *J Clin Densitom* 2006;9:66–71. [PubMed: 16731433]

20. Chapurlat RD, Duboeuf F, Marion-Audibert HO, Kalpakcioglu B, Mitlak BH, Delmas PD. Effectiveness of instant vertebral assessment to detect prevalent vertebral fracture. *Osteoporos Int* 2006;17:1189–1195. [PubMed: 16758133]
21. Gilsanz V, Loro ML, Roe TF, Sayre J, Gilsanz R, Schulz EE. Vertebral size in elderly women with osteoporosis. Mechanical implications and relationship to fractures. *J Clin Invest* 1995;95:2332–2337. [PubMed: 7738196]
22. Takada M, Wu CY, Lang TF, Genant HK. Vertebral fracture assessment using the lateral scoutview of computed tomography in comparison with radiographs. *Osteoporos Int* 1998;8:197–203. [PubMed: 9797902]
23. Schousboe JT, Debold CR. Reliability and accuracy of vertebral fracture assessment with densitometry compared to radiography in clinical practice. *Osteoporos Int* 2006;17:281–289. [PubMed: 16172798]
24. Jiang G, Eastell R, Barrington NA, Ferrar L. Comparison of methods for the visual identification of prevalent vertebral fracture in osteoporosis. *Osteoporos Int* 2004;15:887–896. [PubMed: 15071725]
25. Szulc P, Munoz F, Marchand F, Sornay-Rendu E, Delmas PD. Similar prevalence of vertebral fractures despite different approaches to define reference data. *Bone* 2003;32:441–448. [PubMed: 12689689]
26. Grados F, Roux C, de Vernejoul MC, Utard G, Sebert JL, Fardellone P. Comparison of four morphometric definitions and a semiquantitative consensus reading for assessing prevalent vertebral fractures. *Osteoporos Int* 2001;12:716–722. [PubMed: 11605736]
27. Rea JA, Chen MB, Li J, Blake GM, Steiger P, Genant HK, Fogelman I. Morphometric X-ray absorptiometry and morphometric radiography of the spine: a comparison of prevalent vertebral deformity identification. *J Bone Miner Res* 2000;15:564–574. [PubMed: 10750572]
28. Leidig-Bruckner G, Genant HK, Minne HW, Storm T, Thamsborg G, Bruckner T, Sauer P, Schilling T, Soerensen OH, Ziegler R. Comparison of a semiquantitative and a quantitative method for assessing vertebral fractures in osteoporosis. *Osteoporos Int* 1994;4:154–161. [PubMed: 8069055]
29. Moselewski F, O'Donnell CJ, Achenbach S, Ferencik M, Massaro J, Nguyen A, Cury RC, Abbara S, Jang IK, Brady TJ, Hoffmann U. Calcium concentration of individual coronary calcified plaques as measured by multidetector row computed tomography. *Circulation* 2005;111:3236–3241. [PubMed: 15956133]
30. Dawber TR, Meadors GF, Moore FE. Epidemiological approaches to heart disease: the Framingham Study. *Am J Publ Health* 1951;41:279–286.
31. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med* 1975;4:518–525. [PubMed: 1208363]
32. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, D'Agostino RB Sr, Fox CS, Larson MG, Murabito JM, O'Donnell CJ, Vasani RS, Wolf PA, Levy D. The Third Generation Cohort Of The National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* 2007;165:1328–1335. [PubMed: 17372189]
33. Pou KM, Massaro JM, Hoffmann U, Lieb K, Vasani RS, O'Donnell CJ, Fox CS. Patterns of abdominal fat distribution: the Framingham Heart Study. *Diab Care* 2009;32:481–485.
34. Genant HK, Wu CY, van Kuijk C, Nevitt MC. Vertebral fracture assessment using a semiquantitative technique. *J Bone Miner Res* 1993;8:1137–1148. [PubMed: 8237484]
35. Hoffmann U, Massaro JM, Fox CS, Manders E, O'Donnell CJ. Defining normal distributions of coronary artery calcium in women and men (from the Framingham Heart Study). *Am J Cardiol* 2008;102:1136–1141. e1131. 1141. [PubMed: 18940279]
36. Lenchik L, Rogers LF, Delmas PD, Genant HK. Diagnosis of osteoporotic vertebral fractures: importance of recognition and description by radiologists. *AJR Am J Roentgenol* 2004;183:949–958. [PubMed: 15385286]
37. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
38. Fleiss, J. Statistical method for rates and proportions. Wiley; New York: 1981. The measurement of interrater agreement; p. 211–236.

39. Hospers IC, van der Laan JG, Zeebregts CJ, Nieboer P, Wolffenbuttel BH, Dierckx RA, Kreeftenberg HG, Jager PL, Slart RH. Vertebral fracture assessment in supine position: comparison by using conventional semiquantitative radiography and visual radiography. *Radiology* 2009;251:822–828. [PubMed: 19380691]
40. Zmuda JM, Cauley JA, Glynn NW, Finkelstein JS. Posterior-anterior and lateral dual-energy X-ray absorptiometry for the assessment of vertebral osteoporosis and bone loss among older men. *J Bone Miner Res* 2000;15:1417–1424. [PubMed: 10893693]
41. Olmez N, Kaya T, Gunaydin R, Vidinli BD, Erdogan N, Memis A. Intra- and interobserver variability of Kleerekoper's method in vertebral fracture assessment. *Clin Rheumatol* 2005;24:215–218. [PubMed: 15940554]
42. Kleerekoper, MPA.; Ellis, BI. Measurement of vertebral fracture rates in osteoporosis. In: Christiansen, C.; Nordin, BEC.; Parfitt, AM.; Peck, WA.; Riggs, BL., editors. *Copenhagen International Symposium on Osteoporosis*. Department of Clinical Chemistry, Glostrup Hospital; Copenhagen: 1984. p. 103-109.
43. Bauer JS, Muller D, Ambekar A, Dobritz M, Matsuura M, Eckstein F, Rummeny EJ, Link TM. Detection of osteoporotic vertebral fractures using multidetector CT. *Osteoporos Int* 2006;17:608–615. [PubMed: 16437195]
44. Muller D, Bauer JS, Zeile M, Rummeny EJ, Link TM. Significance of sagittal reformations in routine thoracic and abdominal multislice CT studies for detecting osteoporotic fractures and other spine abnormalities. *Eur Radiol* 2008;18:1696–1702. [PubMed: 18418607]
45. Williams AL, Al-Busaidi A, Sparrow PJ, Adams JE, Whitehouse RW. Under-reporting of osteoporotic vertebral fractures on computed tomography. *Eur J Radiol* 2009;69:179–183. [PubMed: 17913429]
46. Woo EK, Mansoubi H, Alyas F. Incidental vertebral fractures on multidetector CT images of the chest: prevalence and recognition. *Clin Radiol* 2008;63:160–164. [PubMed: 18194691]

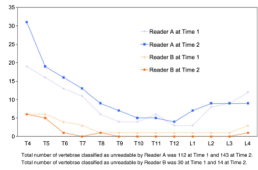


Fig. 1. Number of unreadable vertebrae by vertebral level independently assessed on CT lateral scout views by two readers and two time points for 100 individuals, Framingham CT Study

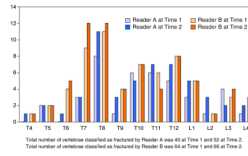


Fig. 2. Number of fractures according to vertebral level independently assessed on CT lateral scout views by two readers and two time points for 100 individuals, Framingham CT Study

Table 1

Characteristics of sex- and age-stratified convenience sample for reliability study of vertebral fracture assessment of 100 participants in the Framingham Offspring and Third Generation Multidetector Computed Tomography Study, 2002-2005

	Women (N=50) N or mean	Men (N=50) N or mean
Age (years)		
50-54	4	3
55-59	4	3
60-64	5	9
65-69	9	7
70-74	8	11
75+	20	17
Height (cm)	158	173
Weight (kg)	70	85
Body mass index (kg/m ²)	28	28

Table 2

Number of individuals according to number of vertebral fractures T4 to L4 independently assessed on CT lateral scout views by two radiologists and two time points, Framingham CT Study

	Participants (N=100)					
	Reader A		Reader B			
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
	N	N	N	N	N	N
Number of fractures						
0	76	77	65	65	65	65
1	15	9	19	19	17	17
2	5	6	7	7	10	10
≥3	4	8	8	9	8	8
Individuals with ≥1 fracture	24	23	35	35	35	35

Table 3

Semi-quantitative grade of prevalent vertebral fracture (T4-L4) independently assessed on CT lateral scout views, per vertebra and per person, by two readers at two time points, Framingham CT Study

	Reader A		Reader B	
	Time 1 N (%)	Time 2 N (%)	Time 1 N (%)	Time 2 N (%)
Semi-quantitative fracture grade per vertebra, N=1,300				
Grade 0 (normal)	1,260 (97)	1,248 (96)	1,236 (95)	1,234 (95)
Grade 1 (mild)	25 (2)	38 (3)	41 (3)	44 (3)
Grade 2 (moderate)	15 (1)	13 (1)	15 (1)	16 (1)
Grade 3 (severe)	0 (-)	1 (<1)	8 (<1)	6 (<1)
Total vertebrae grade ≥ 1 (mild-severe)	40 (3)	52 (4)	64 (5)	66 (5)
Semi-quantitative fracture grade per person ^a , N=100				
Grade 0 (normal)	76	77	65	65
Grade 1 (mild)	13	13	22	23
Grade 2 (moderate)	11	9	6	6
Grade 3 (severe)	0	1	7	6
Total persons grade ≥ 1 (mild-severe)	24	23	35	35

^aIndividuals with more than one fracture are classified according to most severe fracture grade

Table 4

Agreement for grade ≥ 1 and grade ≥ 2 prevalent fractures, per vertebra, for two readers at two time points, Framingham CT Study

Vertebrae (N=1300)					
		Intra-reader agreement between time 1 and time 2		Inter-reader agreement between reader A and reader B	
		Reader A	Reader B	Time 1	Time 2
Fracture definition					
\geq Grade 1 (mild-severe)					
Kappa	0.55	0.77	0.56	0.59	0.59
95% CI	0.43-0.67	0.69-0.85	0.44-0.68	0.48-0.70	0.48-0.70
\geq Grade 2 (moderate-severe)					
Kappa	0.76	0.98	0.68	0.72	0.72
95% CI	0.58-0.93	0.93-1.00	0.51-0.85	0.55-0.89	0.55-0.89

Table 5

Agreement for grade ≥ 1 prevalent fractures, per vertebra, according to spinal region, for two readers at two time points, Framingham CT Study

	<u>Intra-reader agreement between time 1 and time 2</u>		<u>Inter-reader agreement between reader A and reader B</u>	
	Reader A	Reader B	Time 1	Time 2
T4-T9				
Kappa	0.38	0.73	0.43	0.43
95% CI	0.17-0.59	0.61-0.85	0.24-0.61	0.27-0.60
T10-T12				
Kappa	0.65	0.79	0.61	0.72
95% CI	0.46-0.83	0.64-0.93	0.42-0.79	0.55-0.88
L1-L4				
Kappa	0.66	0.87	0.76	0.78
95% CI	0.42-0.90	0.72-1.00	0.55-0.96	0.59-0.97